

Mimicking Intermolecular Interactions of Tight Protein-Protein Complexes for Small-Molecule Antagonists

David Xu^{2,3†}, Dr. Khuchtumur Bum-Erdene^{1†}, Dr. Yubing Si², Dr. Donghui Zhou¹,
Mona Ghozayel¹, and Dr. Samy Meroueh^{1,2*}

¹Department of Biochemistry and Molecular Biology, ²Center for Computational Biology and Bioinformatics, Indiana University School of Medicine, Indianapolis, Indiana, 46202

³Department of BioHealth Informatics, Indiana University School of Informatics and Computing, Indianapolis, Indiana, 46202

*Corresponding author

†Authors contributed equally

Corresponding Author: Samy Meroueh

Department of Biochemistry and Molecular Biology

Indiana University School of Medicine

410 W. 10th Street, HITS 5000

Indianapolis, IN 46202

Tel: (317) 274-8315

Fax: (317) 278-9217

E-mail: smeroueh@iu.edu

ABSTRACT

Tight protein-protein interactions ($K_d < 100$ nM) that occur over a large binding interface ($> 1,000$ Å²) are highly challenging to disrupt with small molecules. Historically, the design of small molecules to inhibit protein-protein interactions has focused on mimicking the position of interface protein ligand side chains. Here, we explore mimicry of the pairwise intermolecular interactions of the native protein ligand with residues of the protein receptor to enrich commercial libraries for small-molecule inhibitors of tight protein-protein interactions. We use the high-affinity interaction ($K_d = 1$ nM) between the urokinase receptor (uPAR) and its ligand urokinase (uPA) to test our methods. We introduce three methods for rank-ordering small molecules docked to uPAR: (i) a new fingerprint approach that represents uPA's pairwise interaction energies with uPAR residues; (ii) a pharmacophore approach to identify small molecules that mimic the position of uPA interface residues; and (iii) a combined fingerprint and pharmacophore approach. Our work led to small molecules with novel chemotypes that inhibited a tight uPAR•uPA protein-protein interaction with single-digit micromolar IC₅₀s. We also report the extensive work that identified several of the hits as either lacking stability, thiol reactive, or redox active. This work suggests that mimicking the binding profile of the native ligand and the position of interface residues can be an effective strategy to enrich commercial libraries for small-molecule inhibitors of tight protein-protein interactions.

INTRODUCTION

Protein-protein interactions range from weak ($K_d > 1000$ nM), moderate (100 nM $< K_d < 1000$ nM), to tight ($K_d < 100$ nM).^[1-3] They are often classified as primary, secondary, or tertiary.^[4] Primary interfaces are generally simple, involving a short linear peptide bound to the surface of another protein. Secondary interactions consist of an α -helix or β -turn that binds to a well-defined cavity of the receptor. Tertiary interactions are more complex, featuring large binding interfaces that often involve several secondary structures such as α -helices and β -strands. Kastiris and co-workers found that 68% of a set of 144 curated co-crystallized protein-protein interactions were both tight and occurred over a large binding interface (> 1000 Å²).^[5] Yet, despite the gradual increase in the number of small-molecule protein-protein interaction inhibitors,^[6-13] only a handful among them are inhibitors of tight protein-protein interactions. Small-molecule inhibitors of tight interactions tend to be much larger than typical drugs, and generally have poor ligand binding efficiencies, which could explain the tendency for these compounds to fail in clinical trials. The development of small molecules that disrupt tight protein-protein interactions could expand the number of druggable proteins for the development of therapeutic agents.

Considering the ever-expanding size of commercial compound libraries, virtual screening could provide an avenue for developing chemical starting points that can be turned into potent inhibitors of tight protein-protein interactions. To the best of our knowledge, one study has used virtual screening to identify small-molecule inhibitors of a tight protein-protein interaction without extensive cycles of design and synthesis.^[14] The most common approach for discovery of protein-protein inhibitors involves the identification of a weak-affinity fragment whose affinity is optimized by growing the fragment into neighboring pockets. For Bcl-xL•Bax^[10] and IL-2•IL2-R α ,^[15] fragment-based approaches and synthesis of derivatives to optimize binding to pockets at the protein-protein interfaces led to highly potent small-molecule inhibitors of the protein-protein interactions. Recently, virtual screening of fragment libraries led to hit compounds and were

optimized into potent inhibitors of KEAP1•NRF2.^[6] In the case of uPAR•uPA, we used computational screening of commercial databases that led to IPR-803, which binds to uPAR with sub-micromolar binding affinity and disrupts the protein-protein interaction with single-digit micromolar IC₅₀s.^[16, 17] The compound was discovered by virtual screening against a set of uPAR structures sampled from molecular dynamics simulations. Computational studies that include molecular dynamics simulations predicted a binding mode for the compound along with a critical salt-bridge interaction with an Arg-53^[18] residue. Our predicted binding mode and interaction with Arg-53 was recently independently confirmed by a crystal structure of an analog of IPR-803 bound to uPAR.^[19]

Historically, most rational approaches for the design of small-molecule inhibitors of protein-protein interactions have focused on mimicking the position of amino acids located on the protein ligand of a protein-protein interaction.^[20-22] Several studies have used interface residues of the protein ligand of a protein-protein interaction to guide the design of small-molecule inhibitors in virtual screening and lead optimization.^[23-29] The most common approach is based on pharmacophore modeling to enrich libraries for compounds that possessed substituents that not only adopted the same position as the amino acid side chain, but also possessed similar physicochemical properties to the side chain. This strategy has worked reasonably well, although it is worth mentioning that there are no examples to date of small molecules that disrupt tight protein-protein interactions that emerged directly from virtual screening. Another strategy consists of finding molecules that bind directly to the receptor with the hope that these compounds will disrupt the protein-protein interaction. This strategy has never led to inhibitors of tight protein-protein interactions. This is attributed to the fact that mere binding to the receptor is not sufficient and critical residues, sometimes referred to as hot spots, must be engaged.

Here, we explore a new approach that is focused on mimicking the pairwise binding profile of the native protein ligand to the protein receptor of a tight protein-protein interaction. We use a

fingerprint approach to represent the binding profile of the protein ligand. We use this fingerprint to screen commercial chemical libraries for small molecules that mimic the pairwise interaction profile of the native protein ligand. We also consider the strategy of combining the fingerprint approach to the standard pharmacophore method that identifies molecules that mimic the position of protein ligand amino acids. We use the tight uPAR•uPA protein-protein interaction as a platform to test these methods. We dock a library of commercially-available compounds to uPAR and rank compounds using the binding profile of the native ligand following three different methods. Compounds are tested for activity using fluorescence polarization and microtiter-based ELISA confirm disruption of the uPAR•uPA interaction. We also test for direct binding with microscale thermophoresis. All active hits are tested for thiol reactivity, redox activity, and stability. An analog-by-catalog procedure to explore structure-activity relationships led to the selection and testing of several derivatives for each hit compound. To the best of our knowledge, this is the first example of the use of structure-based virtual screening that leads to small-molecule inhibitors of a tight protein-protein interaction.

RESULTS

uPAR•uPA as a Platform to Test Rank-Ordering Methods. The uPAR•uPA interface consists primarily of a β -turn on the protein ligand uPA ensconced in a large pocket on the protein receptor uPAR, leading to an interaction that is both tight ($K_d = 1$ nM) and stable ($k_{off} = 10^{-4}$ s $^{-1}$) (Fig. 1 and Fig. S1a). Critical residues exist on both uPAR and uPA. In a comprehensive alanine scanning study, 15 residues on uPAR resulted in a significant decrease in binding affinity ($\Delta\Delta G \geq 1$ kcal mol $^{-1}$) (Fig. S1b and Table S1).^[30] Many of these residues, including Leu-55, Tyr-57, Leu-66, and Leu-150, are located within the binding pocket of uPAR. Recently, using molecular dynamics (MD) simulations followed by free energy calculations, we found residues on uPAR that are engaged in strong van der Waals and electrostatic interactions with uPA, but do not result in a significant decrease in binding affinity; these include Arg-53 and Thr-127 (Fig. S1c and Table

S2).^[17] On uPA, the side chain of five residues extend into the hydrophobic pocket of uPAR: Lys-23, Tyr-24, Phe-25, Ile-28, and Trp-30 (Fig. S1a).^[31]

A New Fingerprint Method to Rank-Order Compounds Based on their Ability to Mimic the Binding Profile of uPA to Residues on uPAR. Although previous studies have used key residues on the protein ligand to guide the design of small-molecule inhibitors, interactions with residues on the receptor have been generally ignored. Here, we utilize the interaction energies in the native protein-protein interaction to select top candidates that emerge from virtual screening of chemical libraries. We introduce a new approach that uses the native ligand to identify compounds that mimic the protein ligand's interaction with key receptor interface residues. To accomplish this, we use a fingerprint method summarized in Fig. 1. These fingerprints consist of strings of bits with length equal to the number of residues on the protein target, in our case uPAR. Each bit in the fingerprint corresponds to the interaction energy between the compound and a residue on uPAR. If the interaction energy between the ligand and the residue is greater than a threshold, the value of the bit is assigned to '1'. For compounds, the interaction energy consists of the computational decomposition energy. A value of '1' is assigned to a bit if the total decomposition energy ($\Delta E_{\text{Residue}}$) is less than $-1.0 \text{ kcal mol}^{-1}$. For the native protein ligand, uPA, we generate two types of fingerprints based on either experimental data or computational decomposition energy. The first type of fingerprint is constructed using the experimentally-determined alanine scanning data of the uPAR•uPA complex (Fig. S1b and Table S1). In this fingerprint, a value of '1' is assigned to a bit if the change in free energy following mutation of the residue to alanine ($\Delta\Delta G_{\text{AlaScan}}$) is greater than $1.0 \text{ kcal mol}^{-1}$. The second fingerprint is constructed using the decomposition energies from the molecular dynamics simulation of the uPAR•uPA complex (Fig. S1c and Table S2). A value of '1' is assigned to a bit if the total decomposition energy ($\Delta E_{\text{Residue}}$) is less than $-1.0 \text{ kcal mol}^{-1}$. If the threshold is not met, the value of the bit is '0'.

Following the docking of small molecules from commercial libraries to uPAR, a fingerprint is generated for each protein-compound structure. Each of these compound fingerprints is compared to the native protein ligand uPA fingerprint. Compounds with the most similar fingerprints to the protein ligand uPA are given higher priority. We use the Tanimoto distance (T_d) to compare the similarity between compounds and protein ligand uPA fingerprint (Fig. 1). T_d is defined as the ratio between the number of bits in the fingerprint where both uPA and the compound have a value of '1' over the number of '1' bits in the uPA fingerprint. The fingerprint generated from either alanine scanning or energy decomposition only includes positions where the corresponding uPA fingerprint has a value of '1'. However, the Tanimoto distance does not consider the positions of the specific bits when used to rank-order compounds. Similarly, the limited length of each fingerprint results in compounds sharing similar Tanimoto distances. If the Tanimoto distance of two compounds is equal, the total enthalpy from the MM-GBSA calculation of the compound (ΔE_{GBTOT}) is used to give higher priority to the compound with higher predicted binding affinity.

Application of the Fingerprint Method to Rank-Order Compounds using uPAR Interface Residues. We use the uPAR•uPA interaction as a platform to test our fingerprint method to rank-order compounds based on their interaction with receptor interface residues. The positions of the residues at the uPAR•uPA interface that were used to generate fingerprints are shown in Fig. 2a. Four residues on uPAR are present in both the uPA fingerprints based on the experimental alanine scanning and the fingerprints from energy decomposition: Leu-55, Leu-66, Leu-150, and His-166.

We separately rank-order the 5.1 million docked compounds based on their T_d value using (i) the uPA alanine scanning fingerprint and (ii) the uPA decomposition energy fingerprint. We select the top 500 candidates from each type of fingerprint. We examined how these compounds bind to each of these residues on uPAR (Fig. 2b). For compounds identified using the uPA

fingerprint derived from energy decomposition, over 90% of the selected compounds interact favorably with Arg-53, Leu-55, Leu-66, Leu-150, and Ala-255. His-166 and Asp-25 interact with 41% and 86% of the compounds, respectively. For compounds identified using the uPA fingerprint derived from the experimental alanine scanning experimental data, 95% of the selected compounds interact with Leu-55, Leu-66, Leu-150, and His-166. Only 36% interacted with Arg-53, a residue that was included in the decomposition but not the alanine scanning fingerprint.

The top 500 compounds from both the energy decomposition and alanine scanning search strategies were independently clustered to 50 compounds using hierarchical clustering. Among the 50 compounds from each strategy, 29 from the uPA alanine scanning fingerprint and 24 from the uPA decomposition energy fingerprint were purchased for experimental validation. The fingerprints and MM-GBSA scores of the 29 and 24 compounds from the alanine scanning and decomposition energy fingerprints are listed in Table S3 and Table S4, respectively. These 53 compounds were initially tested for binding to uPAR using a fluorescence polarization (FP) assay that we have previously developed (Fig. 2c).^[16] The assay consists of a fluorescently labeled α -helical peptide (AE147-FAM) that binds to uPAR at the uPAR•uPA interface. One compound, **1** (IPR-2797), inhibited by more than 40%. A concentration-dependent study led to a K_i of 7.1 ± 1.2 μ M (Fig. 2d). A follow-up study using a microtiter ELISA method to analyze the compound inhibition of the uPAR•uPA_{ATF} interaction was performed. The compound did not show activity in the ELISA even at 100 μ M. We next assessed **1** for both reactivity and stability. The potential for **1** to covalently react with cysteine residues of a protein was evaluated using a (*E*)-2-(4-mercaptostyryl)-1,3,3-trimethyl-3H-indol-1-ium (MSTI)-based assay.^[32] The compound did not react with MSTI suggesting that it is not thiol reactive (Fig. S2). The compound was tested for redox activity by a Horseradish Peroxidase-Phenol Red (HRP-PR) assay and was found to be redox inactive at 100 μ M concentration (Fig. S3). Compound stability was tested in methanol, phosphate-buffered saline (PBS), and in the presence of uPAR by high-performance liquid

chromatography-mass spectrometry (HPLC-MS). The compound showed the same retention time in HPLC and the mass remained the same, indicating that the compound is stable (Fig. S4).

The predicted binding mode of **1** shows that the benzene of the benzofuran moiety overlaps with Phe-25 on uPA (Fig. 3a). In addition, a nitrogen in the piperazine ring of the compound is located near the positively charged amine on the side chain of Lys-23. Starting with the structure of **1**, we searched commercially-available libraries for analogs to conduct a preliminary structure-activity relationship (SAR) study. A set of 36 derivatives of **1** were identified, purchased, and screened at 50 μ M using our FP assay (Fig. 3b). Three compounds, **2** (IPR-2944), **3** (IPR-2962), and **4** (IPR-2966), showed 75%, 66%, and 82% inhibition, respectively (Fig. 3c). The compounds inhibited the uPAR•AE147-FAM interaction in a concentration-dependent manner (Fig. S5), although all were weaker than the parent compound. The most potent derivative, **4** ($K_i = 8.6 \pm 1.1$ μ M), contains a fluorine atom on the aromatic ring of its benzofuran moiety.

Selecting Rank-Ordered Compounds using uPA Interface Residues. We explore another ranking method that strictly uses the interface residues located on the ligand protein to guide the selection of compounds. Our hypothesis is that small molecules docked to uPAR that mimic the side chain position of interface residues on the protein ligand will disrupt the uPAR•uPA interaction. At the uPAR•uPA interface, the side chain of five interface residues on uPA extend into the hydrophobic pocket of uPAR: Lys-23, Tyr-24, Phe-25, Ile-28, and Trp-30 (Fig. 4a). We use a pharmacophore approach^[33, 34] to identify compounds with substituents that occupy the same position as the side chains of these residues. This approach consists of searching for small molecules that possess substituents that overlap with similar moieties (pharmacophores) on the side chain of amino acids. For example, a compound that possesses a benzene group that occupies the same position as the aromatic ring (pharmacophore) of a tyrosine residue is expected to disrupt binding of the residue to uPAR. The features of the pharmacophore model

are shown in Fig. 4a. In the pharmacophore model, the ϵ -amine on Lys-23 was modeled using a positive charge, while the benzene rings of Tyr-24 and Phe-25 were modeled using aromatic rings. We assigned separate aromatic ring features to the benzene and pyrrole rings on the indole of Trp-30.

For each of the 5.1 million docked compounds to uPAR, we determined whether there was an overlap with the defined pharmacophores on uPA. This resulted in 21312, 809846, 1297014, and 23047 matches for Lys-23, Tyr-24, Phe-25, and Trp-30, respectively. We identified 1899 compounds that overlapped with 3 of these 4 residues, and no compounds that overlapped with all four residues. These compounds were hierarchically clustered to 200 using atom triplet Daylight fingerprints. We identified 130 commercially-available compounds that were purchased. These compounds were tested for binding to uPAR using our FP assay (Fig. 4b). In total, approximately 1.8 million of the 5.1 million docked compounds overlapped with at least one of the pharmacophores corresponding to an interface residue on uPA (Fig. 4c). Among compounds that overlapped with a single residue, 54% and 27% matched the pharmacophores of Phe-25 and Tyr-24, respectively. Less than 2% of compounds overlapped with either the Lys-23 or Trp-30 pharmacophore. In contrast, 16% of compounds overlapped with both the Tyr-24 and Phe-25 pharmacophores.

We initially tested the 130 compounds for activity using our FP assay at 50 μ M. We selected six compounds (**5-10**) that inhibited more than 40% (Fig. 4b). Compounds **5** (IPR-2477), **6** (IPR-2496), and **9** (IPR-2532) overlapped with Tyr-24, Phe-25, and Trp-30, while **7** (IPR-2500), **8** (IPR-2529), and **10** (IPR-2565) overlapped with Lys-23, Tyr-24, and Phe-25 (Fig. 4d). A concentration-dependent study for these compounds led to K_i values that ranged from 6 to 97 μ M (Fig. S6a). A follow-up study using a microtiter ELISA to analyze the compound inhibition of the uPAR•uPA_{ATF} interaction was performed (Fig. S6b). Although the ELISA cannot be used to obtain inhibition constants, it is a useful assay to determine whether compounds bind and disrupt the

protein-protein interaction between uPA and uPAR. Four compounds, namely **6**, **7**, **9**, and **10** showed activity in the ELISA assay with IC_{50} s ranging from 7 to 230 μ M.

We assessed the reactivity and stability of all hit compounds. Compound **6** contains a Betti base that may cause the compound to be unstable and reactive, while **10** was thought to have a potential activated thiol group. The MSTI thiol reactivity assay was performed for each of the hits (Fig. S2). Compound **10** readily reacted with MSTI as evidenced by a decrease in the fluorescence of MSTI. Compound **6** displayed no detectable MSTI reactivity, but we suspected that this was due to the unstable nature of the compound. The HRP-PR redox activity assay showed no significant redox capacity for compounds **5-10** (Fig. S3). At this point, we decided to pursue compounds **5**, **6**, **8**, and **9**. HPLC-MS stability assay for **8** and **9** showed the same single peak for both methanol and PBS buffers, indicating that the compounds were stable (Fig. S4).

A set of 55 derivatives of **5** were purchased and screened using the FP assay at 50 μ M. A number of compounds displayed significant inhibition (Fig. S7a). Three compounds (Fig. S7b), **11** (IPR-2603), **12** (IPR-2605), and **13** (IPR-2606), were chosen to be further studied at several concentrations using FP (Fig. S7c) and ELISA (Fig. S7d). Compound **11** displayed better K_i than the parent in the FP assay while all three compounds showed activity in the ELISA. MSTI assay of these three derivatives detected no thiol reactivity (Fig. S7e), but the reaction might be masked by assay interference, as all three compounds are yellow when dissolved in the assay buffer. The compounds showed no appreciable redox activity (Fig. S3). However, the HPLC-MS study revealed that all the compounds were mixtures as shown for **12** (Fig. S4). This suggests that the compound series may be unstable and inhibits through a non-specific mechanism. This is further supported by the fact that despite the differences in structure between **11**, **12**, and **13**, the three compounds exhibit similar K_i s in the FP assay and IC_{50} in the ELISA.

Compound **6** was pursued with reservation considering the reactive Betti base. Compounds **14-17** were purchased, and we synthesized **18** (IPR-2804) that lacked the hydroxyl

group of the Betti base (Fig. S8a). Compounds were initially tested at 50 μ M and then in a concentration-dependent manner using the FP assay and ELISA (Fig. S8b). The compounds with the Betti base moiety inhibited substantially, while **18** showed no activity in either assay. The complete lack of activity of **18**, despite its structural similarity with the parent compound, strongly suggested that the activity of **6** and derivatives was non-specific. This is further confirmed by the MSTI assay, which shows **18** to be non-reactive while **14-17** are reactive (Fig. S8c). Compound **17** (IPR-2665) did not bind to uPAR using microscale thermophoresis (data not shown). To determine whether the compounds reacted with uPA_{ATF} as well as uPAR, the ELISA setup was slightly modified, whereby we pre-incubated immobilized uPA_{ATF} with the compounds for 30 min, followed by the washing step, which discarded all compounds in the well, followed by addition of uPAR. Despite the complete lack of contact of the compound with uPAR, both compounds **6** and **17** displayed significant inhibition at 50 μ M concentration, while our reversible uPAR inhibitor, IPR-1109,^[35] was washed away and did not inhibit (Fig. S8d). This suggested that the compounds were also forming adducts with uPA further supporting a non-specific mechanism. To test for specificity, **17** was tested in two other FP assays against TEAD4 transcription factor and voltage-gated calcium channel β -subunit 3 (Ca ν β ₃) proteins; they inhibited in both assays (Fig. S8e). Compounds **6** and **17** showed no redox activity on the HRP-PR assay (Fig. S3). HPLC-MS analysis of **6** showed that the compound broke down in methanol and PBS buffer, resulting in weak UV signals. This confirmed that the compounds were unstable. The total ion count chromatogram showed three distinct peaks, the intact **6** and two fragments (Fig. S8f). This compound series was not pursued further.

Compound **8** binds to uPAR with double-digit micromolar affinity ($K_i = 37.2 \pm 19.8 \mu$ M) based on the FP assay. The binding mode from the virtual screen shows that the compound's morpholino, methylphenyl, and methoxyphenyl groups overlap with Lys-23, Tyr-24, and Phe-25, respectively (Fig. 5a). A search of the ZINC chemical library identified 37 derivatives that were

purchased and screened at 50 μM using our FP assay (Fig. 5b). Two derivatives, **19** (IPR-2916) and **20** (IPR-2922) (Fig. 5c) were chosen and tested by concentration-dependent FP (Fig. S9a) and ELISA (Fig. S9b). The compounds showed weak K_i and IC_{50} values, but they provide a starting point for the development of more potent inhibitors of the very tight protein-protein interaction.

Compound **9** binds to uPAR with affinity near 100 μM (FP $K_i = 97.4 \pm 61.2 \mu\text{M}$, ELISA $\text{IC}_{50} = 230.6 \pm 41.7 \mu\text{M}$). The virtual screening binding mode of **9** is shown in Fig. 5d. On one end of the compound, the dimethoxybenzene overlaps with Trp-30. On the other end, the acyl morpholine is buried into the binding pocket. A benzene ring near this moiety overlaps with Phe-25, while a pyrimidine in the core overlaps with Tyr-24. A total of 8 derivatives were purchased and screened by the single-concentration FP assay (Fig. 5e). None of the derivatives showed strong similarity to the parent compound. As a result, most of these derivatives showed even weaker binding and inhibition than **9**. The best compounds were **21** (IPR-2940) and **22** (IPR-2942) (Fig. 5f), which we tested by concentration-dependent FP (Fig. S10a) and ELISA (Fig. S10b).

Selecting Rank-Ordered Compounds using both uPA and uPAR Interface Residues.

We wondered whether combining our fingerprint method with the pharmacophore approach could yield small-molecule uPAR•uPA inhibitors. We combined the two search methods to identify a set of 69 compounds that overlapped with three of the four interface residues on uPA as well as engage residues in the uPA binding pocket on uPAR. A set of 39 compounds selected from among the 69 compounds were purchased for binding studies. The fingerprints and MM-GBSA scores of the 15 and 24 compounds from the alanine scanning and decomposition energy fingerprints are listed in Table S5 and Table S6, respectively. The 39 compounds were tested in our FP assay (Fig. 6a). Seven compounds (**23-29**) were assessed using a concentration-dependent manner to determine the K_i values (Fig. 6b and S11a). Compounds **23** (IPR-2986), **26**

(IPR-2992), **27** (IPR-2993), **28** (IPR-3089), and **29** (IPR-3193) overlapped with Lys-23, Tyr-24, and Phe-25 on uPA, while **24** (IPR-2987) and **25** (IPR-2989) overlapped with Tyr-24, Phe-25, and Trp-30. In comparison to hits that emerged from fingerprint or pharmacophore methods, the compounds had better K_i values that ranged from 6 to 52 μM . Only **25**, **26**, **28**, and **29** inhibited uPAR•uPA_{ATF} based on a concentration-dependent study using our ELISA (Fig. S11b). The compounds had IC_{50} values of 68.3 ± 11.0 , 140.6 ± 19.0 , 172.8 ± 42.5 , and 24.8 ± 2.2 μM in the ELISA, respectively. The presence of α,β -unsaturated carbonyls on **23**, **24** and **27** suggested potential reactivity with residue on uPAR and uPA. However, none of the hits from this screen showed reactivity with the activated thiol of the MSTI compound suggesting that the activity of the compounds is unlikely due to covalent bond formation (Fig. S2). Compounds **24** and **27** displayed slight redox capacity in the HRP-PR assay, while compounds **23**, **25**, **26**, **28**, and **29** showed no redox activity (Fig. S3). We focused our attention on **26**, **28**, and **29** as these compounds do not contain any problematic moiety and showed no covalent reaction or redox activity. HPLC-MS analysis of these three compounds (Fig. S4) showed the compounds to be stable in both methanol and PBS buffers, with similar retention times.

Compound **26** binds into the uPAR•uPA pocket, mimicking the side chains of uPA and engaging interface residues on uPAR (Fig. 7a). A benzene moiety overlaps with Phe-25 on uPA. In the binding mode, a morpholino group is located between Lys-23 and Tyr-24 of uPA. In addition, a methyl substituent on the core quinoline ring points towards Trp-30. The core structure of **26** was used to identify derivatives. The derivatives we identified showed modifications at five substituents on **26** (Fig. 7b). A set of 136 derivatives of **26** were purchased and screened at 50 μM (Fig. 7c). The best hits were tested in concentration-dependent manner and their K_i s ranged from 2 to 37 μM (Table 1 and Fig. S12a). The compounds were further tested in the uPAR•uPA_{ATF} ELISA assay to determine whether they can inhibit the protein-protein interaction (Table 1 and Fig. S12b). Only **32** (IPR-3026) and **39** (IPR-3116) failed to inhibit in the ELISA. The best

derivative among the **26** derivatives was **30** (IPR-3011). The binding mode of **30** shows that the additional moiety fits into a pocket lined by Asn-157, His-166, Leu-168, and Ala-255 on uPAR (Fig. 7a). The FP and ELISA inhibition curves of **26** and **30** are shown in Fig. 7d and Fig. 7e, respectively. The K_i and IC_{50} in the FP and ELISA assays for **30** are 2.5 ± 0.3 and 15.5 ± 1.4 μ M, respectively. Compounds **26** (Fig. 7f) and **30** (Fig. 7g) were tested using microscale thermophoresis to assess direct binding to uPAR. The resulting K_d of **26** and **30** towards uPAR were 5.8 ± 1.3 and 2.0 ± 0.4 μ M, respectively, consistent with the FP data for these compounds. Compound **30** and several of the other **26** derivatives have limited solubility. Like its parent **26**, **30** displayed no significant redox activity at 100 μ M (Fig. S3).

The binding mode of **28** shows overlap with Lys-23, Tyr-24, and Phe-25 on uPA (Fig. 8a). A set of 59 derivatives of **28** were purchased and screened by single concentration FP at 50 μ M (Fig. 8b). The six best compounds, **44-49** (Fig. 8c), were tested in a concentration-dependent manner using the FP assay (Fig. S13). While the analogs had K_i values ranging from 5 to 160 μ M, the compounds did not inhibit in the ELISA. The first set of analogs, **44-46**, modify both the biphenyl and methyl group on the benzimidazole of the parent compound, yielding compounds that were less potent than the parent compound. The second set of analogs, **47-49**, modifies only the methyl group on the benzimidazole. Compound **49** (IPR-3485) lacks the methyl group entirely, resulting in threefold weaker inhibition constant than the parent. Compounds **47** and **48** possess aromatic substituents instead of the methyl group, resulting in K_i s of 5 μ M in the FP assay. The lack of inhibition in the ELISA suggests that the compounds, while robust, may not be engaging the right set of residues to disrupt the full protein-protein interaction.

The binding mode of **29** reveals overlap with Lys-23, Tyr-24, and Phe-25 on uPA (Fig. 8d). A set of 20 derivatives of **29** were purchased and screened at 50 μ M (Fig. 8e). The five best compounds, **50-54** (Fig. 8f), were tested at multiple concentrations using FP (Fig. S14a) and uPAR•uPA_{ATF} ELISA (Fig. S14b). All but **53** (IPR-3235) (FP $K_i = 5.6 \pm 0.6$ μ M, ELISA $IC_{50} = 52.0$

$\pm 4.1 \mu\text{M}$) and **54** (IPR-3236) (FP showed no inhibition; ELISA $\text{IC}_{50} = 65.3 \pm 12.3 \mu\text{M}$) had limited solubility in the two assays. Compound **29** was tested for direct binding in the MST assay and binds to uPAR with a K_d of $22.7 \pm 11.5 \mu\text{M}$ (Fig. S14c).

DISCUSSION

The design of small-molecule inhibitors of protein-protein interactions has primarily focused on developing small molecules that bear substituents that mimic the position of amino acid side chains of the protein ligand in a protein-protein interaction. Here, we complement this approach by exploring a strategy that searches for small molecules that mimic the binding profile of the native protein ligand to the receptor of a protein-protein interaction. We hypothesize that small molecules that mimic the interaction of the native protein ligand to receptor are more likely to disrupt tight protein-protein interactions. To test this hypothesis, we introduced a quantitative approach to enable the comparison of the binding profiles of compounds to the binding profile of the protein ligand. We use a bitwise fingerprint to represent the pairwise interactions with amino acids on the receptor. When the ligand (protein or compound) engages a residue above a threshold, we assign the bit as '1'. The pairwise binding is based on the decomposition energy method that was introduced by Gohlke and co-workers to study the effect of individual amino acids on a protein-protein interaction.^[36] The decomposition energy consists of the intermolecular energy between the ligand and each amino acid. This energy includes van der Waals, electrostatic, and polar and non-polar solvation energies. Following the creation of a fingerprint for the native protein ligand, in our case uPA, we used structure-based virtual screening to identify small molecules that shared a similar fingerprint. To do this, we docked a large number of compounds to uPAR, and generated fingerprints for all these compounds using the predicted binding pose. Compounds were ranked based on how closely their fingerprint matched the native ligand's (e.g. uPA). To accomplish this, we borrow from the cheminformatics field and use the Tanimoto distance to quantify the similarity between fingerprints. The top candidates that

emerged from this screen were purchased and tested for binding to uPAR. Among them, we found one hit, **1** (IPR-2797), that inhibited a fluorescently-labeled peptide with an inhibition constant K_i of 7 μ M. Analogs of **1** were purchased, confirming the activity of the parent compound and providing an opportunity for structure-activity relationships. Neither **1** nor its derivatives inhibited in our ELISA, which includes the entire protein-protein interaction interface. This suggests that improvements could be made to this method, such as perhaps focusing only on the most critical residues on the receptor. A more stringent threshold for picking interface amino acids could make this possible. Another possibility is that there exists a combination of residues that must be engaged to disrupt a protein-protein interaction, and that compound **1** does not engage the right combination of residues. Regardless, the compound offers an excellent starting point to develop potent small-molecule inhibitors of the tight uPAR•uPA protein-protein interaction. We found the compound to have good solubility, does not react with uPAR, and is not redox active. It is also stable in methanol and buffer as evidenced by LC-MS analysis.

Another strategy that we followed was driven by the hypothesis that small molecules that possess substituents that mimic the position of protein ligand amino acid side chains will more likely disrupt the protein-protein interaction. This is a widely used method, but to the best of our knowledge has never been applied for the discovery of small-molecule inhibitors of tight protein-protein interactions. We resorted to pharmacophore modeling to score compounds based on how effectively they overlap with native protein ligand residue side chains of uPA. This approach resulted in more hit compounds than the screen using receptor amino acids alone. However, close inspection of the structure of these compounds revealed potentially problematic groups, such as a Betti base in compound **6** (IPR-2496) that could result in unstable compounds, or a thiol reactive moiety in compound **10** (IPR-2565) that may lead to adduct formation with nucleophilic residues on uPAR. Among all the compounds, we confirmed that **10** is thiol reactive. Compound **6** was unstable, as expected, despite the single-digit inhibition in both the FP and

ELISA assays for a series of derivatives. In fact, we confirmed that the activity of the compound was due to non-specific reactivity through the synthesis of **18** (IPR-2804), a derivative that lacked the Betti base. Despite the lack of obvious unstable or reactive moieties for compound **5** (IPR-2477), we found it to be unstable and its activity is likely due to assay interference or reactivity with assay or protein. Compounds **8** (IPR-2529) and **9** (IPR-2532) were the most robust compounds we identified. Only compound **9** inhibited in both the FP and ELISA suggesting that it could be a good starting point for the development of uPAR•uPA inhibitors. Its large size, however, may make it difficult to optimize. Compound **8** did not inhibit uPAR•uPA in our ELISA suggesting that the compound binds, but it may not effectively mimic uPA residues. It is also possible that the compound binds to residues on uPAR that negate the benefits of mimicking uPA.

Finally, we wondered whether the use of interface residues on both uPAR and uPA could lead to better inhibitors from virtual screening. We combined our fingerprint and pharmacophore methods to rank-order compounds docked to uPAR. It is interesting that this method led to even more hit compounds than using fingerprint or pharmacophore alone. A total of seven hits were identified. Despite the initial concern that three of the compounds had potentially problematic α,β -unsaturated carbonyls, such as in **23** (IPR-2986), **24** (IPR-2987), or **27** (IPR-2993), none of the compounds were found to be thiol reactive. One compound had an acylhydrazine moiety that could also be unstable at low pH, although our work is done at pH 7 suggesting that the compound should be stable. An interesting feature of these compounds compared to those that emerged from using strictly the pharmacophore method is that they had fewer rotatable bonds overall, and two compounds, namely **26** (IPR-2992) and **28** (IPR-3089) were fragment-like. Compound **26** was particularly interesting as it inhibited, albeit weakly, in both our FP and ELISA assays. Starting with **26**, we followed an analog-by-catalog approach and purchased several derivatives. Among the derivatives, we discovered several compounds, including **30** (IPR-3011), which exhibited substantially higher binding affinity than the parent fragment-like compound. We

confirmed direct binding of both **26** and **30** using microscale thermophoresis with K_d values that were similar to the K_i s values measured by FP. **30** also possessed substantially better IC_{50} s (single-digit micromolar range) in the disruption of the full uPAR•uPA interaction. Future optimization will focus on improving solubility of these derivative compounds, and exploring additional substituents for **26**. The methyl group located on the quinoline ring points towards the side chain of a critical tryptophan on uPA. The introduction of moieties that mimic the tryptophan side chain may result in substantially greater potency.

In sum, we present a new approach to identify small-molecule inhibitors of tight protein-protein interactions that uses the native ligand's pairwise intermolecular interactions with the receptor of a protein-protein interaction. When combined with a pharmacophore approach that uses the native protein ligand interface amino acids, we identified robust small-molecule inhibitors of the tight uPAR•uPA. To the best of our knowledge, this is the first example of a virtual screen that uses the crystal structure of a tight protein-protein interaction and identified single-digit micromolar small-molecule inhibitors. These results suggest that while commercial libraries do not cover chemical space that is typical of protein-protein interaction inhibitors, it is possible to identify robust starting points that could be used to develop small-molecule inhibitors of tight protein-protein interactions. The results also show that virtual screening is also prone to nuisance compounds as several of the small molecules that initially showed promising activity were working through a non-specific mechanism. Finally, it is worth mentioning that small-molecule inhibitors that emerged from this work are structurally distinct from inhibitors that we previously identified for uPAR. We compared the structure of compounds **1-54** to our previously reported uPAR•uPA inhibitors IPR-803^[16] and IPR-1110.^[17, 35] The similarity between these compounds was assessed using atom triplet Daylight fingerprints. We find that generally, the Tanimoto similarity between compounds **1-54** and our previously described inhibitors range from 0.05 to 0.10.

MATERIALS AND METHODS

Virtual Screening. A set of commercially-available compounds from ChemDiv Inc. (San Diego, CA), ChemBridge Corporation (San Diego, CA), Life Chemicals (Munich, Germany), Princeton BioMolecular Research Inc. (Princeton, NJ), Specs (Zoetermeer, Netherlands), and Vitas-M Laboratory Ltd (Hong Kong) were retrieved from ZINC.^[37] Small molecules in this library possessing pan-assay interference compound (PAINS)^[38] or rapid elimination of swill (REOS)^[39] moieties were filtered out using the Canvas package in Schrödinger (Schrödinger LLC, New York, NY, 2015). This resulted in a compound library of approximately 5.1 million small molecules. Individual MOL2 formatted files were converted to PDBQT format using the *prepare_ligand4.py* script in MGLTools.^[40]

The structure of the uPAR•uPA complex (PDB ID: 3BT1) was retrieved and prepared using Protein Preparation Wizard in Schrödinger.^[41] Bond orders were assigned, hydrogen atoms were added, and disulfide bonds were created. Vitronectin (chain B) was removed and the missing loop at residues Arg-83 and Ala-84 were introduced using the Prime module. The resulting structure was protonated at pH 7.0 using PROPKA^[42] and separated into its respective monomeric chains. The uPAR structure (chain U) was converted to PDBQT format using the *prepare_receptor4.py* script in MGLTools.

The compound library was docked to the prepared uPAR structure using AutoDock Vina.^[40] The binding pocket was centered at the uPAR•uPA interface with a box with dimensions of 21 Å × 21 Å × 21 Å. All other parameters were set to default values. The docked conformations were converted back to MOL2 format using in-house Python scripts for additional analysis.

uPAR Interface Residues. To find compounds that overlapped with residues on uPAR in the uPAR•uPA complex, we resorted to a fingerprint approach that utilizes interaction energies between the receptor and ligand. The workflow of the fingerprint approach is described in Fig. S15. We determined the interaction energies of each docked compound to individual residues of uPAR using the Generalized Born Surface Area (GBSA) method in the Amber14 and

AmberTools15 software packages.^[43] Each docked compound was assigned Gasteiger charges and gaff^[44] atom types using the *antechamber* program.^[45] Additional force field parameters were generated using the *parmchk* program. Topology and coordinate files for the docked complex and individual receptor and ligand were generated with ff14SB^[46] and gaff^[44] force fields using the *tLeap* program. These topology and coordinate files were used as inputs to calculate the free energies and per-residue decomposition energies in the *MMPBSA.py* script.^[47] The *MMPBSA.py* script was modified to include the missing atom radius for iodine atoms.^[48] The calculation using the Generalized Born (GB) method was performed with *sander* and Onufriev's GB model.^[49, 50] Solvent-accessible surface area (SASA) calculations were switched to the icosahedron (ICOSA) method, where surface areas are computed by recursively approximating a sphere around an atom, starting from an icosahedron. Salt concentration was set to 0.1 M. Compounds with combined internal and solvation terms (ΔE_{GBTOT}) greater than -5.0 kcal mol⁻¹ were discarded.

For each docked compound, we generate a one-dimensional array with length equal to the total number of residues of the uPAR structure. In this vector, each position corresponds to an individual residue of uPAR. Each position is assigned a value of '1' (ON) or '0' (OFF) based on the residue decomposition energy at that position and acts as a fingerprint for that compound. If the energy at the given residue is less than -1.0 kcal mol⁻¹, we assign the position a value of '1'. Otherwise, we assign the position a value of '0'. Residues for the fingerprints were identified from two sources: (i) an experimentally-determined alanine scanning of uPAR from Gårdsvoll and coworkers^[30] (Table S1); and (ii) a previously described molecular dynamics (MD) simulation of the uPAR•uPA complex (Table S2).^[17] Similar to the construction of the compound-specific bitwise arrays, we create vectors for each type of fingerprint where each position corresponds to an interaction energy of the uPAR•uPA complex. In the vector corresponding to the experimental alanine scan, a position was assigned a value of '1' if the $\Delta\Delta G$ at that residue is greater than 1.0 kcal mol⁻¹ and '0' otherwise. In the vector corresponding to the per-residue decomposition

energies, a position is assigned a value of '1' if the total energy (ΔE_{GBTOT}) at that residue is less than -1.0 kcal mol⁻¹ and '0' otherwise.

In both fingerprints, only a small portion of uPAR will have values of '1' with its native ligand uPA. Therefore, we reduce the length of each fingerprint to only include positions with '1' bits in the uPAR•uPA complex. For each docked compound, we calculate the Tanimoto distance between the fingerprints of the complex and the compound in a bitwise manner. The fingerprint of the uPAR•uPA complex consists of only '1' bits. Thus, this distance can be simply calculated by summing the number of '1' bits in the compound fingerprint and dividing by the length of the fingerprint. Compounds were rank-ordered based on their Tanimoto distance, and in cases where compounds had the same Tanimoto distance, we used ΔE_{GBTOT} to rank these compounds.

uPA Interface Residues. A pharmacophore-based approach was used to identify docked compounds that overlapped with and mimicked interface residues on uPA. We used four residues of uPA at the uPAR•uPA interface: Lys-23, Tyr-24, Phe-25, and Trp-30. For each residue, we defined a pharmacophore hypothesis corresponding to the physiochemical properties of the individual residue's side chain using the Phase package in Schrödinger.^[33, 34] Phase has six built-in types of pharmacophore features: (i) hydrogen bond acceptor, (ii) hydrogen bond donor, (iii) hydrophobe, (iv) negative ionizable, (v) positive ionizable, and (vi) aromatic ring. We assigned a positive charged feature to the ϵ -amine on Lys-23 and aromatic rings features to the aromatic rings of Tyr-24, Phe-25, and Trp-30. A single pharmacophore feature was assigned to the benzene rings of Tyr-24 and Phe-25, while two separate pharmacophores were assigned to the pyrrole and benzene rings of the bicyclic indole on Trp-30. We searched for compounds containing ligand moieties that matched a corresponding pharmacophore feature. A compound that matched either of the two aromatic pharmacophore features on Trp-30 was considered to overlap and mimic the residue. Each pharmacophore feature was screened independently of one another. The aromatic and positively charged pharmacophores are represented as spheres

centered on the side chain moiety with radii of 1.5 Å and 0.75 Å, respectively. Compound conformers from virtual screening were used to identify matches without refinement using Phase's default fitness function. In this scoring function, three factors are used to describe the degree to which a compound matches a pharmacophore feature: (i) the site score, which describes how well the compound superimposes the pharmacophore feature, (ii) the vector score, which describes the cosine angle between the normal vector of an aromatic ring on the compound with an aromatic feature, and (iii) the volume score, which describes the proportion of the total volume of the pharmacophore feature overlapped by the compound. Compounds with either root-mean-square deviation (RMSD) overlap greater than 1.2 Å or did not overlap with the pharmacophore feature were discarded. For the aromatic pharmacophores, no consideration was given to the angle between the normal vectors of an aromatic feature and the orientation of an aromatic ring. All other parameters were set at default values. The remaining compounds that matched a given pharmacophore was retained without sorting compounds by Phase's internal scoring function. Compounds that matched 3 of the 4 residues were retained.

Selection of Compounds. The top-ranking compounds following virtual screening using uPAR and uPA residues were retrieved and clustered using the Canvas package in Schrödinger. A hashed binary fingerprint corresponding to atom triplets of Daylight invariant atom types were generated for these top-ranking compounds. Compounds were then hierarchically clustered using their atom triplet fingerprints and average linkage clustering. The Tanimoto similarity between a pair of fingerprints was used as the distance metric. Compounds corresponding to the cluster centers from hierarchical clustering were purchased for experimental validation.

Microtiter-Based ELISA for uPAR•uPA. uPAR without the glycosylphosphatidylinositol (GPI) anchor was obtained by a purification process as previously described.^[51] High-binding microplates (Greiner Bio-One, Kremsmünster, Austria) were incubated for 2 h at 4 °C with 100 µL of 4 µg mL⁻¹ of uPA_{ATF} in PBS for immobilization as previously described.^[16] The plate was

washed with 0.05% Tween-20 in PBS buffer between each step. A 1:1 mixture of Superblock buffer in PBS (Thermo Fisher Scientific, Inc. Waltham, MA) with 0.04 M NaH_2PO_4 and 0.3 M NaCl buffer was used for blocking at room temperature for 45 min. Following removal of the blocking buffer and washing, 100 μL of 0.85 nM uPAR in PBS with 0.01% triton X-100 was added with 100 to 0.4 μM compounds in 1% v/v DMSO. Following incubation for 30 min and subsequent washing steps, biotinylated human uPAR antibody (1:3000 dilution of 0.2 mg mL^{-1} BAF807, R&D Systems, Minneapolis, MN) in PBS containing 1% bovine serum albumin (BSA) was added to the wells (100 μL /well) and incubated for 1 h to allow for the detection of bound uPAR. Following washing, streptavidin-horseradish-peroxidase in PBS containing 1% BSA was added to the wells and incubated for 20 min. The signal developed in the presence of 3,3',5,5'-tetramethylbenzidine (TMB) in phosphate-citrate buffer (pH 5) and hydrogen peroxide was stopped by adding H_2SO_4 solution and was detected using a SpectraMax M5e (Molecular Devices, Sunnyvale, CA). When compounds were insoluble and visible precipitation was observed, the data points at the high concentrations were not included in the calculation of IC_{50} values.

Fluorescence Polarization (FP) Assay. Polarized fluorescence intensities were measured using EnVision Multilabel plate readers (PerkinElmer, Waltham, MA) with excitation and emission wavelengths of 485 and 535 nm, respectively.^[16] Samples were prepared in Thermo Scientific Nunc 384-well black microplate in duplicates. First, the compounds were serially diluted in DMSO and further diluted in 1× PBS buffer with 0.01% Triton X-100 for a final concentration of 200 – 0.2 μM . Triton X-100 was added to the buffer to avoid compound aggregation. 5 μL of the compound solution and 40 μL of PBS with 0.01% Triton X-100 containing uPAR was added to the wells and incubated for at least 15 min to allow the compound to bind to the protein. Finally, 5 μL of fluorescent AE147-FAM peptide solution was added for a total volume of 50 μL in each well resulting in final uPAR and peptide concentrations of 320 nM and 100 nM respectively. The final DMSO concentration was 2% v/v, which had no effect on the binding of the peptide. Controls

included wells containing only the peptide and wells containing both protein and peptide each in duplicates to ensure the validity of the assay. A unit of millipolarization (mP) was used for calculating percentage inhibition of the compounds. When compounds were insoluble and visible precipitation was observed, the data points at the high concentrations were not included in the calculation of IC₅₀ values. Inhibition constants were calculated from the IC₅₀ values using the K_i calculator available at http://sw16.im.med.umich.edu/software/calc_ki/.

Microscale Thermophoresis (MST). uPAR was labeled with NT-495 fluorescent dye (Nanotemper, Munich, Germany) according to the manufacturer's instructions. Compounds were serially diluted in DMSO and further diluted in PBS buffer with 0.025% v/v Tween-20. 10 μ L of fluorescently-labeled uPAR and 10 μ L of compound solution were combined to final concentrations of 40 nM fluorescently-labeled uPAR, 2% v/v DMSO and compound concentrations ranging from 200 μ M to 0.1 μ M. The protein-compound solution was incubated for 10 min at room temperature in the dark, before being taken up in Monolith NT.115 series standard-treated capillaries. The capillaries were measured on Monolith NT.115 (Nanotemper, Munich, Germany) at 25 °C, with LED power at 40% and MST power at 40%, and the MST was measured for 30 s. The data was analyzed using the "Thermophoresis with T-jump" function within the NanoTemper Affinity Analysis version 2.0.2 software (Nanotemper, Munich, Germany). The data was then fit with the " K_d Model" function within the software to calculate the K_d .

(E)-2-(4-mercaptostyryl)-1,3,3-trimethyl-3H-indol-1-ium (MSTI) Assay. MSTI assay was performed according to the manufacturer's recommendations (Kerafast, Inc. Boston, MA).^[32] A 10 mM solution of acetyl-MSTI was added to 50 mM PBS buffered solution at pH 12.0 with 50% v/v methanol in a ratio of 1:10 v/v. After mixing and incubating for 2 min at room temperature, the solution was diluted with 50 mM PBS at pH 7.4, containing 0.01% NP-40 and 5% v/v methanol to generate a final concentration of 30 μ M MSTI at pH 7.4. 19.6 μ L aliquots of the MSTI solution was dispensed in 384-well flat-bottom black plate and 0.4 μ L of 5 mM compounds in DMSO were

added make a 100 μM final concentration. Unreacted MSTI solution without added compound, but with equal amount of DMSO, was used as a negative control, while unactivated acetyl-MSTI solution with DMSO was used as a positive control. The plate was then incubated with shaking for 30 min at room temperature and the fluorescence intensities were measured using a Flexstation 3 plate reader (Molecular Devices, Sunnyvale, CA) at excitation and emission wavelengths of 510 and 650 nm, respectively.

Horseradish Peroxidase-Phenol Red (HRP-PR) Redox Activity Assay. HRP-PR assay was performed according to the published protocol.^[52, 53] In brief, 20 μL of 300 μM compounds in 3% v/v DMSO in Hank's balanced salt solution (HBSS) (Cat. No. SH30268.02; HyClone, Logan, UT) was dispensed into a 384-well clear, flat-bottomed polystyrene plate (Cat. No. 781101; Greiner Bio-One, Monroe, NC). Controls with 3% v/v DMSO and 300 μM H_2O_2 were dispensed. 20 μL of 2.4 mM fresh dithiothreitol (DTT) in HBSS was added to each well. For the H_2O_2 controls, 20 μL HBSS with no DTT was added. After 5 min incubation at RT, 20 μL solution of 300 $\mu\text{g ml}^{-1}$ Phenol Red (Cat. No. P-2417; Sigma-Aldrich, St. Louis, MO), 180 $\mu\text{g ml}^{-1}$ HRP (Cat. No. P-2088; Sigma-Aldrich, St. Louis, MO) was added. After 20 min incubation at room temperature, the absorbance was read at 610 nm on a SpectraMax M5e (Molecular Devices, Sunnyvale, CA).

High-Performance Liquid Chromatography-Mass Spectrometry (HPLC-MS). Compounds at 200 μM were incubated in methanol, PBS, or 30 μM uPAR in PBS for 1 h at room temperature. The samples were injected onto a Kinetex 2.6 μm XB-C18 100 Å column (Cat. No. 00B-4496-E0; Phenomenox, Torrance, CA) on an Agilent 6130 Quadrupole LC/MS system (Agilent, Santa Clara, CA). Compounds **6** (IPR-2496), **12** (IPR-2605), and **29** (IPR-3193) were eluted by a linear gradient from Buffer A (H_2O) to Buffer B (acetonitrile, 5 mM NH_4OAc) over 15 min. Compounds **1** (IPR-2797), **8** (IPR-2529), **9** (IPR-2532), **26** (IPR-2992), and **28** (IPR-3089) were eluted by a linear gradient from Buffer A (H_2O , 0.1% formic acid) to Buffer B (acetonitrile,

0.1% formic acid) over 15 min. Column elution was tracked by UV absorption at 256 nm and the masses were detected by positive ion mode.

ABBREVIATIONS USED

Ca _v β ₃	voltage-dependent L-type calcium channel subunit beta-3
FP	fluorescence polarization
GB	Generalized Born
HRP-RP	Horseradish Peroxidase-Phenol Red
K _d	disassociation constant
MM-GBSA	molecular mechanics generalized Born surface area approach
MST	microscale thermophoresis
MSTI	(<i>E</i>)-2-(4-mercaptostyryl)-1,3,3-trimethyl-3H-indol-1-ium
PAINS	pan-assay interference compound
PPI	protein-protein interaction
REOS	rapid elimination of swill
SAR	structure-activity relationship
TEAD4	transcriptional enhancer factor TEF-3
uPAR	urokinase plasminogen activator surface receptor
uPA	urokinase-type plasminogen activator

ACKNOWLEDGEMENTS

The research was supported by the National Institutes of Health (R01 CA197928) (SOM), the American Cancer Society Research Scholar Grant RSG-12-092-01-CDD (SOM), and the Vera Bradley Foundation (KB). Computer time on the Big Red II and Karst supercomputer at Indiana University is supported in part by Lilly Endowment, Inc., through its support for the Indiana University Pervasive Technology Institute, and in part by the Indiana METACyt Initiative. This research was done using resources provided by the Open Science Grid, which is supported by the National Science Foundation award 1148698, and the U.S. Department of Energy's Office of Science.

REFERENCES

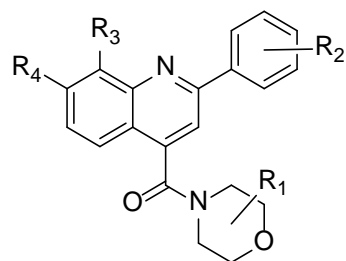
- [1] R. P. Bahadur, P. Chakrabarti, F. Rodier, J. Janin, *J. Mol. Biol.* **2004**, 336(4), 943-955.
- [2] J. R. Perkins, I. Diboun, B. H. Dessailly, J. G. Lees, C. Orengo, *Structure* **2010**, 18(10), 1233-1243.
- [3] M. C. Smith, J. E. Gestwicki, *Expert Rev. Mol. Med.* **2012**, 14, e16.
- [4] M. R. Arkin, Y. Tang, J. A. Wells, *Chem. Biol.* **2014**, 21(9), 1102-1114.
- [5] P. L. Kastiris, I. H. Moal, H. Hwang, Z. Weng, P. A. Bates, A. M. Bonvin, J. Janin, *Protein Sci.* **2011**, 20(3), 482-491.
- [6] T. G. Davies, W. E. Wixted, J. E. Coyle, C. Griffiths-Jones, K. Hearn, R. McMenamin, D. Norton, S. J. Rich, C. Richardson, G. Saxty, H. M. Willems, A. J. Woolford, J. E. Cottom, J. P. Kou, J. G. Yonchuk, H. G. Feldser, Y. Sanchez, J. P. Foley, B. J. Bolognese, G. Logan, P. L. Podolin, H. Yan, J. F. Callahan, T. D. Heightman, J. K. Kerns, *J. Med. Chem.* **2016**, 59(8), 3991-4006.
- [7] C. Tovar, B. Graves, K. Packman, Z. Filipovic, B. Higgins, M. Xia, C. Tardell, R. Garrido, E. Lee, K. Kolinsky, K. H. To, M. Linn, F. Podlaski, P. Wovkulich, B. Vu, L. T. Vassilev, *Cancer Res.* **2013**, 73(8), 2587-2597.
- [8] Q. Ding, Z. Zhang, J. J. Liu, N. Jiang, J. Zhang, T. M. Ross, X. J. Chu, D. Bartkovitz, F. Podlaski, C. Janson, C. Tovar, Z. M. Filipovic, B. Higgins, K. Glenn, K. Packman, L. T. Vassilev, B. Graves, *J. Med. Chem.* **2013**, 56(14), 5979-5983.
- [9] F. Christ, A. Voet, A. Marchand, S. Nicolet, B. A. Desimmie, D. Marchand, D. Bardiot, N. J. Van der Veken, B. Van Remoortel, S. V. Strelkov, M. De Maeyer, P. Chaltin, Z. Debyser, *Nat. Chem. Biol.* **2010**, 6(6), 442-448.
- [10] T. Oltersdorf, S. W. Elmore, A. R. Shoemaker, R. C. Armstrong, D. J. Augeri, B. A. Belli, M. Bruncko, T. L. Deckwerth, J. Dinges, P. J. Hajduk, M. K. Joseph, S. Kitada, S. J. Korsmeyer, A. R. Kunzer, A. Letai, C. Li, M. J. Mitten, D. G. Nettesheim, S. Ng, P. M. Nimmer, J. M. O'Connor,

- A. Oleksijew, A. M. Petros, J. C. Reed, W. Shen, S. K. Tahir, C. B. Thompson, K. J. Tomaselli, B. Wang, M. D. Wendt, H. Zhang, S. W. Fesik, S. H. Rosenberg, *Nature* **2005**, 435(7042), 677-681.
- [11] L. T. Vassilev, B. T. Vu, B. Graves, D. Carvajal, F. Podlaski, Z. Filipovic, N. Kong, U. Kammlott, C. Lukacs, C. Klein, N. Fotouhi, E. A. Liu, *Science* **2004**, 303(5659), 844-848.
- [12] M. H. Roehrl, S. Kang, J. Aramburu, G. Wagner, A. Rao, P. G. Hogan, *Proc. Natl. Acad. Sci. U. S. A.* **2004**, 101(20), 7554-7559.
- [13] M. Lepourcelet, Y. N. Chen, D. S. France, H. Wang, P. Crews, F. Petersen, C. Bruseo, A. W. Wood, R. A. Shivdasani, *Cancer Cell* **2004**, 5(1), 91-102.
- [14] P. Mukherjee, P. Desai, Y. D. Zhou, M. Avery, *J. Chem. Inf. Model.* **2010**, 50(5), 906-923.
- [15] C. G. Wilson, M. R. Arkin, *Curr. Top. Microbiol. Immunol.* **2011**, 348, 25-59.
- [16] M. Khanna, F. Wang, I. Jo, W. E. Knabe, S. M. Wilson, L. Li, K. Bum-Erdene, J. Li, W. S. G, R. Khanna, S. O. Meroueh, *ACS Chem. Biol.* **2011**, 6(11), 1232-1243.
- [17] D. Liu, D. Xu, M. Liu, W. E. Knabe, C. Yuan, D. Zhou, M. Huang, S. O. Meroueh, *Biochemistry* **2017**, 56(12), 1768-1784.
- [18] Q. Huai, A. P. Mazar, A. Kuo, G. C. Parry, D. E. Shaw, J. Callahan, Y. Li, C. Yuan, C. Bian, L. Chen, B. Furie, B. C. Furie, D. B. Cines, M. Huang, *Science* **2006**, 311(5761), 656-659.
- [19] A. F. Rullo, K. J. Fitzgerald, V. Muthusamy, M. Liu, C. Yuan, M. Huang, M. Kim, A. E. Cho, D. A. Spiegel, *Angew. Chem., Int. Ed. Engl.* **2016**, 55(11), 3642-3646.
- [20] L. R. Hoggard, Y. Zhang, M. Zhang, V. Panic, J. A. Wisniewski, H. Ji, *Journal of the American Chemical Society* **2015**, 137(38), 12249-12260.
- [21] D. Rognan, *MedChemComm* **2015**, 6(1), 51-60.
- [22] D. E. Scott, A. R. Bayly, C. Abell, J. Skidmore, *Nat. Rev. Drug Discovery* **2016**, 15(8), 533-550.
- [23] Y. Lu, Z. Nikolovska-Coleska, X. Fang, W. Gao, S. Shangary, S. Qiu, D. Qin, S. Wang, *J. Med. Chem.* **2006**, 49(13), 3759-3762.

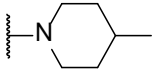
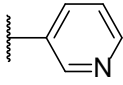
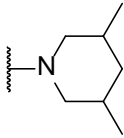
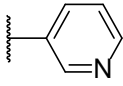
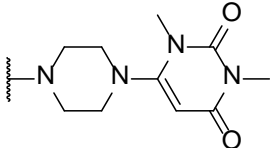
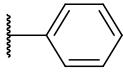
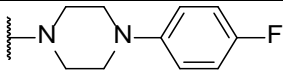
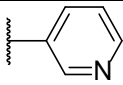
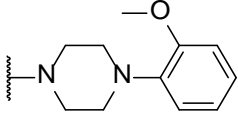
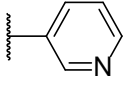
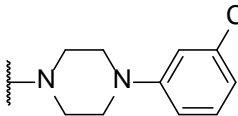
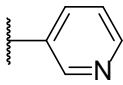
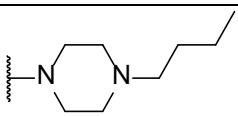
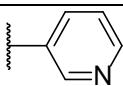
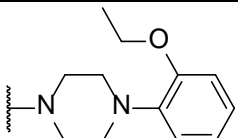
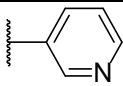
- [24] P. Tortorella, A. Laghezza, M. Durante, I. Gomez-Monterrey, A. Bertamino, P. Campiglia, F. Loiodice, S. Daniele, C. Martini, M. Agamennone, *J. Chem. Inf. Model.* **2016**, 56(6), 1216-1227.
- [25] L. Wang, L. Li, Z. H. Zhou, Z. Y. Jiang, Q. D. You, X. L. Xu, *Eur. J. Med. Chem.* **2017**, 136, 63-73.
- [26] G. Melagraki, E. Ntougkos, V. Rinotas, C. Papaneophytou, G. Leonis, T. Mavromoustakos, G. Kontopidis, E. Douni, A. Afantitis, G. Kollias, *PLoS computational biology* **2017**, 13(4), e1005372.
- [27] F. Gessier, J. Kallen, E. Jacoby, P. Chene, T. Stachyra-Valat, S. Ruetz, S. Jeay, P. Holzer, K. Masuya, P. Furet, *Bioorg. Med. Chem. Lett.* **2015**, 25(17), 3621-3625.
- [28] K. H. Leung, L. J. Liu, S. Lin, L. Lu, H. J. Zhong, D. Susanti, W. Rao, M. Wang, W. I. Che, D. S. Chan, C. H. Leung, P. W. Chan, D. L. Ma, *Methods* **2015**, 71, 38-43.
- [29] H. Li, H. Xiao, L. Lin, D. Jou, V. Kumari, J. Lin, C. Li, *J. Med. Chem.* **2014**, 57(3), 632-641.
- [30] H. Gårdsvoll, B. Gilquin, M. H. Le Du, A. Ménèz, T. J. Jørgensen, M. Ploug, *J. Biol. Chem.* **2006**, 281(28), 19260-19272.
- [31] V. Magdolen, P. Rettenberger, M. Koppitz, L. Goretzki, H. Kessler, U. H. Weidle, B. König, H. Graeff, M. Schmitt, O. Wilhelm, *Eur. J. Biochem.* **1996**, 237(3), 743-751.
- [32] M. M. McCallum, P. Nandhikonda, J. J. Temmer, C. Eyermann, A. Simeonov, A. Jadhav, A. Yasgar, D. Maloney, A. L. Arnold, *J. Biomol. Screening* **2013**, 18(6), 705-713.
- [33] S. L. Dixon, A. M. Smondryev, E. H. Knoll, S. N. Rao, D. E. Shaw, R. A. Friesner, *J. Comput.-Aided Mol. Des.* **2006**, 20(10-11), 647-671.
- [34] S. L. Dixon, A. M. Smondryev, S. N. Rao, *Chem. Biol. Drug Des.* **2006**, 67(5), 370-372.
- [35] D. Liu, D. Zhou, B. Wang, W. E. Knabe, S. O. Meroueh, *ACS Chem. Biol.* **2015**, 10(6), 1521-1534.
- [36] H. Gohlke, C. Kiel, D. A. Case, *J. Mol. Biol.* **2003**, 330(4), 891-913.
- [37] J. J. Irwin, T. Sterling, M. M. Mysinger, E. S. Bolstad, R. G. Coleman, *J. Chem. Inf. Model.* **2012**, 52(7), 1757-1768.

- [38] J. B. Baell, G. A. Holloway, *J. Med. Chem.* **2010**, 53(7), 2719-2740.
- [39] W. P. Walters, A. A. Murcko, M. A. Murcko, *Curr. Opin. Chem. Biol.* **1999**, 3(4), 384-387.
- [40] O. Trott, A. J. Olson, *J. Comput. Chem.* **2010**, 31(2), 455-461.
- [41] G. M. Sastry, M. Adzhigirey, T. Day, R. Annabhimoju, W. Sherman, *J. Comput.-Aided Mol. Des.* **2013**, 27(3), 221-234.
- [42] M. H. M. Olsson, C. R. Søndergaard, M. Rostkowski, J. H. Jensen, *J. Chem. Theory Comput.* **2011**, 7(2), 525-537.
- [43] D. A. Case, J. T. Berryman, R. M. Betz, D. S. Cerutti, I. T.E. Cheatham, T. A. Darden, R. E. Duke, T. J. Giese, H. Gohlke, A. W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T. S. Lee, S. LeGrand, P. Li, T. Luchko, R. Luo, B. Madej, K. M. Merz, G. Monard, P. Needham, H. Nguyen, H. T. Nguyen, I. Omelyan, A. Onufriev, D. R. Roe, A. Roitberg, R. Salomon-Ferrer, C. L. Simmerling, W. Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu, D. M. York, P. A. Kollman, AMBER 2015, University of California, San Francisco, **2015**.
- [44] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* **2004**, 25(9), 1157-1174.
- [45] J. Wang, W. Wang, P. A. Kollman, D. A. Case, *J. Mol. Graphics Modell.* **2006**, 25(2), 247-260.
- [46] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling, *J. Chem. Theory Comput.* **2015**, 11(8), 3696-3713.
- [47] B. R. Miller, 3rd, T. D. McGee, Jr., J. M. Swails, N. Homeyer, H. Gohlke, A. E. Roitberg, *J. Chem. Theory Comput.* **2012**, 8(9), 3314-3321.
- [48] A. Bondi, *J. Phys. Chem.* **1964**, 68(3), 441-451.
- [49] A. Onufriev, D. Bashford, D. A. Case, *Proteins: Struct., Funct., Bioinf.* **2004**, 55(2), 383-394.
- [50] M. Feig, A. Onufriev, M. S. Lee, W. Im, D. A. Case, C. L. Brooks, 3rd, *J. Comput. Chem.* **2004**, 25(2), 265-284.

- [51] B. Jacobsen, H. Gardsvoll, G. Juhl Funch, S. Ostergaard, V. Barkholt, M. Ploug, *Protein Expression Purif.* **2007**, 52(2), 286-296.
- [52] K. M. Soares, N. Blackmon, T. Y. Shun, S. N. Shinde, H. K. Takyi, P. Wipf, J. S. Lazo, P. A. Johnston, *Assay Drug Dev. Technol.* **2010**, 8(2), 152-174.
- [53] P. A. Johnston, K. M. Soares, S. N. Shinde, C. A. Foster, T. Y. Shun, H. K. Takyi, P. Wipf, J. S. Lazo, *Assay Drug Dev. Technol.* **2008**, 6(4), 505-518.

Table 1. Profiles of Analogs of Compound 26 (IPR-2992)

Compound	R ₁	R ₂	R ₃	R ₄	FP K _i (μM) ^[a]	ELISA IC ₅₀ (μM) ^[a]
30 (IPR-3011)			Me	H	2.5 ± 0.3	15.5 ± 1.4
31 (IPR-3015)			Me	H	37.1 ± 1.9	171.5 ± 35.7
32 (IPR-3026)			Me	Cl	15.4 ± 3.4	No inhibition
33 (IPR-3036)			Me	H	5.2 ± 0.5	62.7 ± 10.7
34 (IPR-3037)			Me	H	8.0 ± 1.1	99.5 ± 14.2
35 (IPR-3038)			Me	H	11.9 ± 2.0	67.9 ± 11.2

36 (IPR-3039)			Me	H	8.4 ± 0.6	37.9 ± 2.4
37 (IPR-3040)			Me	H	26.6 ± 7.8	112.2 ± 8.6
38 (IPR-3103)			Me	H	5.2 ± 1.1	31.0 ± 2.1
39 (IPR-3116)			H	H	10.4 ± 3.1	No inhibition
40 (IPR-3117)			H	H	9.2 ± 1.9	331.6 ± 197.0
41 (IPR-3121)			H	H	3.4 ± 0.5	34.4 ± 7.4
42 (IPR-3134)			Me	H	9.9 ± 1.4	122.8 ± 15.3
43 (IPR-3147)			Cl	H	2.5 ± 0.3	26.3 ± 6.8

^[a] Representative of at least two independent experiments, where each concentration point is measured in duplicates (mean \pm SD).

FIGURE LEGENDS

Figure 1. Workflow for the fingerprint method used to identify compounds that mimic the intermolecular binding interactions in the uPAR•uPA complex. The per-residue interaction energies of docked compounds are compared to those of the native protein ligand uPA. These interaction energies are used to generate a bitwise fingerprint, where each position on the fingerprint corresponds to the interaction energy between uPAR and the compound of interest. This fingerprint is compared to fingerprints of the native ligand uPA. Compounds are rank-ordered based on their Tanimoto distance with the fingerprints of uPA and total interaction energy ΔE_{GBTOT} .

Figure 2. A virtual screen utilizing the interface residues of uPAR and validation of hits. a) Residues used in the uPAR fingerprints are colored on the surface of uPAR as follows: (i) Experimental alanine scan (orange), (ii) decomposition (pink), (iii) both (green). uPA is transparently overlaid in cartoon, with the side chain of interface residues in stick. b) Among the top-ranking 500 compounds from each of the fingerprints generated from decomposition energies or experimental alanine scanning, the proportion of compounds that overlap with each fingerprint residue. c) Single-concentration FP screen of compounds resulting from the virtual screen based on uPAR residues. Each compound was screened in duplicate at 50 μM concentration (mean \pm SD). Hit compound **1** (IPR-2797) is highlighted in green. d) Concentration-dependent FP assay measuring the inhibition of uPAR•AE147-FAM peptide interaction by **1** (IPR-2797). Representative of at least two independent experiments, where each concentration point is measured in duplicates (mean \pm SD).

Figure 3. Screening the derivatives of compound **1** (IPR-2797). a) The binding mode of **1** in the uPAR•uPA binding pocket. The compound is shown in yellow. uPAR is shown in white cartoon, with the side chain of interface residues shown in pink stick. uPA is shown in partial transparent cyan cartoon. The side chain of four interface residues on uPA are shown in stick and colored cyan. b) Derivatives of **1** were screened at a single 50 μ M concentration via the uPAR•AE147-FAM peptide FP assay in duplicates (mean \pm SD). Further pursued hits are highlighted in green. c) Chemical structures of the pursued derivative hits.

Figure 4. A virtual screen utilizing four interface residues of uPA and validation of hits. a) Features of the pharmacophore model used to identify compounds that overlap with and mimic the interface residues of uPA. uPAR is shown in the background colored in white and shown in cartoon. uPA is shown in transparent cyan cartoon, with the five interface residues shown in stick. A pharmacophore model was used to assign features to four of the five interface residues (Ile-28 was excluded). The amine on the side chain of Lys-23 was assigned a positive ionizable feature (transparent red circle), while the aromatic rings of Tyr-24, Phe-25, and Trp-30 were assigned aromatic ring features (transparent yellow circles). Two separate pharmacophore features were assigned to each of the two rings on the indole on Trp-30. b) Single-concentration FP screen of compounds resulting from the virtual screen based on uPA interface residues. Each compound was screened in duplicate at 50 μ M concentration (mean \pm SD). Hits that are followed up are highlighted in green while those with problematic moieties are highlighted in red. Chemical structures of the highlighted molecules are shown above. c) Co-occurrence of interface residues among all compounds that overlapped with at least one residues on uPA. d) Overlap between the predicted binding mode of the

hit molecules and the uPA residues are highlighted. FP and microtiter ELISA assays were used to measure the K_i and IC_{50} of the compounds in inhibiting uPAR•AE147-FAM peptide and uPAR•uPA_{ATF} interactions, respectively. Results are based on at least two independent concentration-dependent experiments where each concentration point was measured in duplicates (mean \pm SD).

Figure 5. Screening the derivatives of compounds **8** (IPR-2529) and **9** (IPR-2532). a) The virtual screening binding mode of **8** in the uPAR•uPA binding pocket. The compound is shown in yellow. uPAR is shown in white cartoon, with the side chain of interface residues shown in pink stick. uPA is shown in partial transparent cyan cartoon. The side chain of four interface residues on uPA are shown in stick and colored cyan. b) Derivatives of **8** were screened at a single 50 μ M concentration via FP assay in duplicates (mean \pm SD). Further pursued hits are highlighted in green. c) Chemical structures of the pursued derivative hits of **8**. d) The virtual screening binding mode of **9** (IPR-2532) in the uPAR•uPA binding pocket. The compound is shown in yellow. uPAR is shown in white cartoon, with the side chain of interface residues shown in pink stick. uPA is shown in partial transparent cyan cartoon. The side chain of four interface residues on uPA are shown in stick and colored cyan. e) Derivatives of **9** were screened at a single 50 μ M concentration via FP assay in duplicates (mean \pm SD). Further pursued hits are highlighted in green. f) Chemical structures of the pursued derivative hits of **9**.

Figure 6. A virtual screen utilizing interface residues on both uPAR and uPA. a) Single-concentration FP screen of compounds resulting from the virtual screen based on uPA interface residues. Each compound was screened in duplicate at 50 μ M concentration (mean \pm SD). Hits that are pursued are highlighted in green while those with problematic moieties are highlighted in red. Chemical structures of the

highlighted molecules are shown above. b) Overlap between the predicted binding mode of the hit molecules and the uPA interface residues are highlighted. FP and microtiter ELISA assays were used to measure the K_i and IC_{50} of the compounds in inhibiting uPAR•AE147-FAM peptide and uPAR•uPA_{ATF} interactions, respectively. Results are based on at least two independent concentration-dependent experiments where each concentration point was measured in duplicates (mean \pm SD).

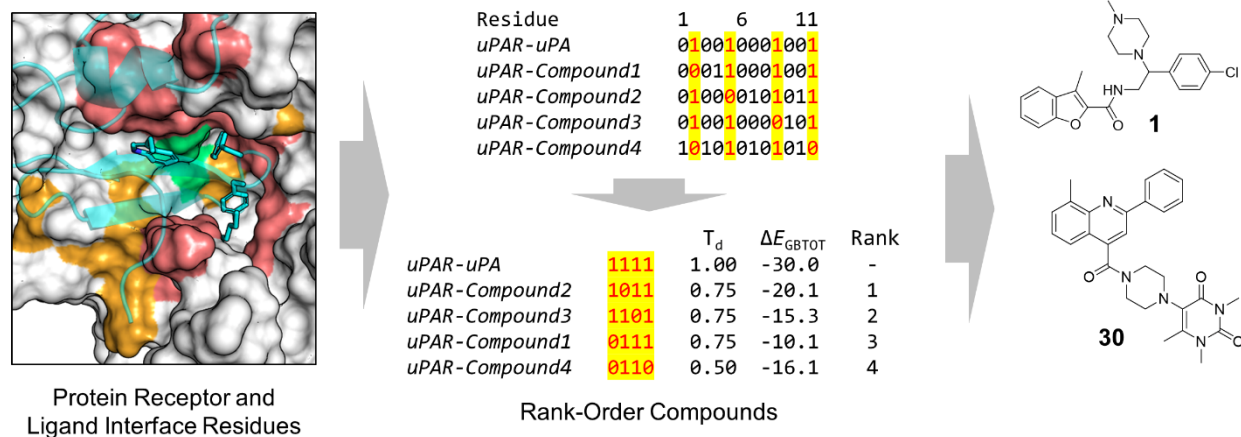
Figure 7. Testing the derivatives of **26** (IPR-2992) leads to **30** (IPR-3011). a) The binding mode of **26** (IPR-2992) and **30** (IPR-3011) in the uPAR•uPA binding pocket. The binding mode of **30** (green) is overlaid on the binding mode of **26** (yellow). The additional ring at R₁ allows **30** to bind deeper in the uPAR•uPA pocket. uPAR is shown in white cartoon, with the side chain of interface residues shown in pink stick. b) The core of **26** was used to identify analogs at 5 positions. Among the analogs discovered was **30** (IPR-3011). c) Derivatives of **26** were screened at a single 50 μ M concentration via the uPAR•AE147-FAM peptide FP assay in duplicates (mean \pm SD). The parent compound **26** is highlighted in orange, while compound **30** is highlighted in green. d) Concentration-dependent FP assay measuring the inhibition of uPAR•AE147-FAM peptide interaction by **26** and **30**. Representative of at least two independent concentration-dependent experiments where each concentration point was measured in duplicates (mean \pm SD). At high concentrations, **30** was insoluble and as such the data points were omitted from curve-fitting. e) Concentration-dependent ELISA assay measuring inhibition of uPAR•uPA_{ATF} interaction by **26** and **30**. Representative of at least two independent concentration-dependent experiments where each concentration point was measured in duplicates (mean \pm SD). At high concentrations, **30** was insoluble and

as such the data points were omitted from curve-fitting. f) MST experiment was performed with 40 nM NT-495-labeled uPAR and varying concentrations of **26**. A representative MST concentration-response curve of the interaction between uPAR and **26** are shown. Two independent experiments were performed in triplicates (mean \pm SD). g) MST experiment was performed with 40 nM NT-495-labeled uPAR and varying concentrations of **30**. A representative MST concentration-response curve of the interaction between uPAR and **30** are shown. Two independent experiments were performed in triplicates (mean \pm SD). At high concentrations, **30** is insoluble and the high concentration points, in light orange, are omitted from curve fitting.

Figure 8. Screening the derivatives of compound **28** (IPR-3089) and **29** (IPR-3193). a) The virtual screening binding mode of **28** in the uPAR•uPA binding pocket. The compound is shown in yellow. uPAR is shown in white cartoon, with the side chain of interface residues shown in pink stick. uPA is shown in partial transparent cyan cartoon. The side chain of four interface residues on uPA are shown in stick and colored cyan. b) Derivatives of **28** were screened at a single 50 μ M concentration via FP assay in duplicates (mean \pm SD). Further pursued hits are highlighted in green. c) Chemical structures of the pursued derivative hits of **28**. K_i values are based on two independent concentration-dependent inhibition assays performed in duplicates (mean \pm SD). d) The virtual screening binding mode of **29** in the uPAR•uPA binding pocket. The compound is shown in yellow. uPAR is shown in white cartoon, with the side chain of interface residues shown in pink stick. uPA is shown in partial transparent cyan cartoon. The side chain of four interface residues on uPA are shown in stick and colored cyan. e) Derivatives of **29** were screened at a single 50 μ M concentration via FP assay in duplicates (mean \pm SD). Further

pursued hits are highlighted in green. f) Chemical structures of the pursued derivative hits of **29**. K_i values are based on two independent concentration-dependent inhibition assays performed in duplicates (mean \pm SD).

Table of Contents.



Mimicking protein-protein interactions: We explore mimicry of the pairwise intermolecular interactions of the native protein ligand with residues of the protein receptor in the tight protein-protein interaction between urokinase and its receptor. We use both fingerprint and pharmacophore approaches to compare the binding profiles of the native ligand and docked compounds. We identify several small molecules with novel chemotypes that inhibited the tight interaction with single-digit micromolar IC_{50} s.

Keywords.

Urokinase receptor

Protein-protein interaction

Small molecule

Virtual screening

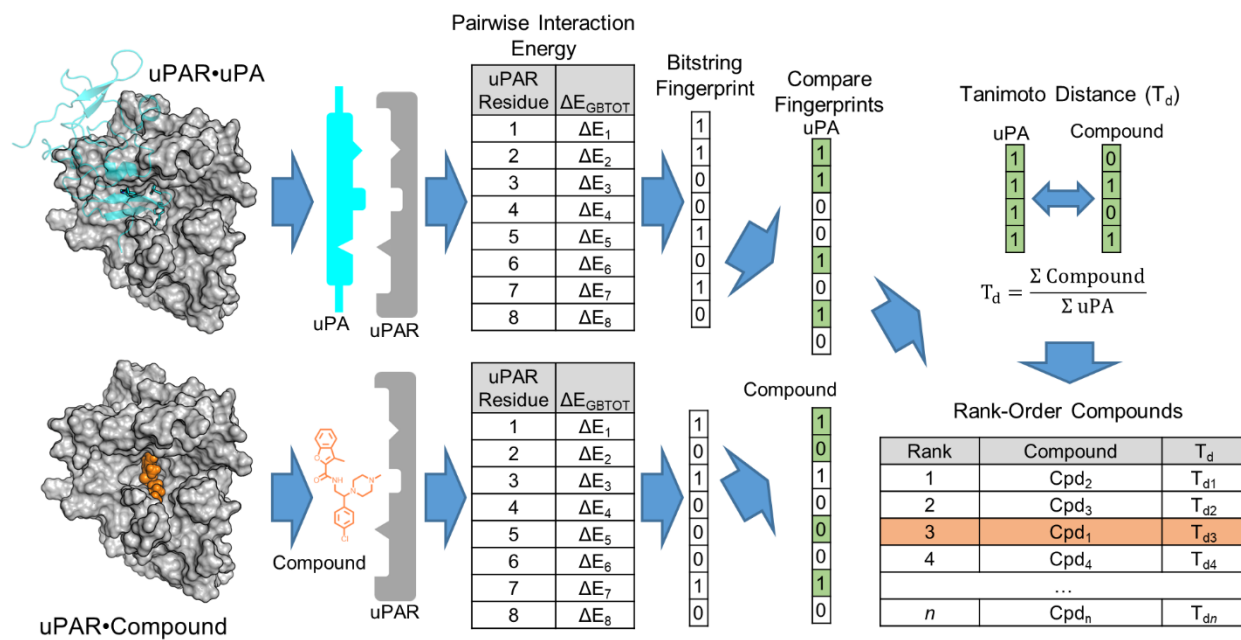


Figure 1

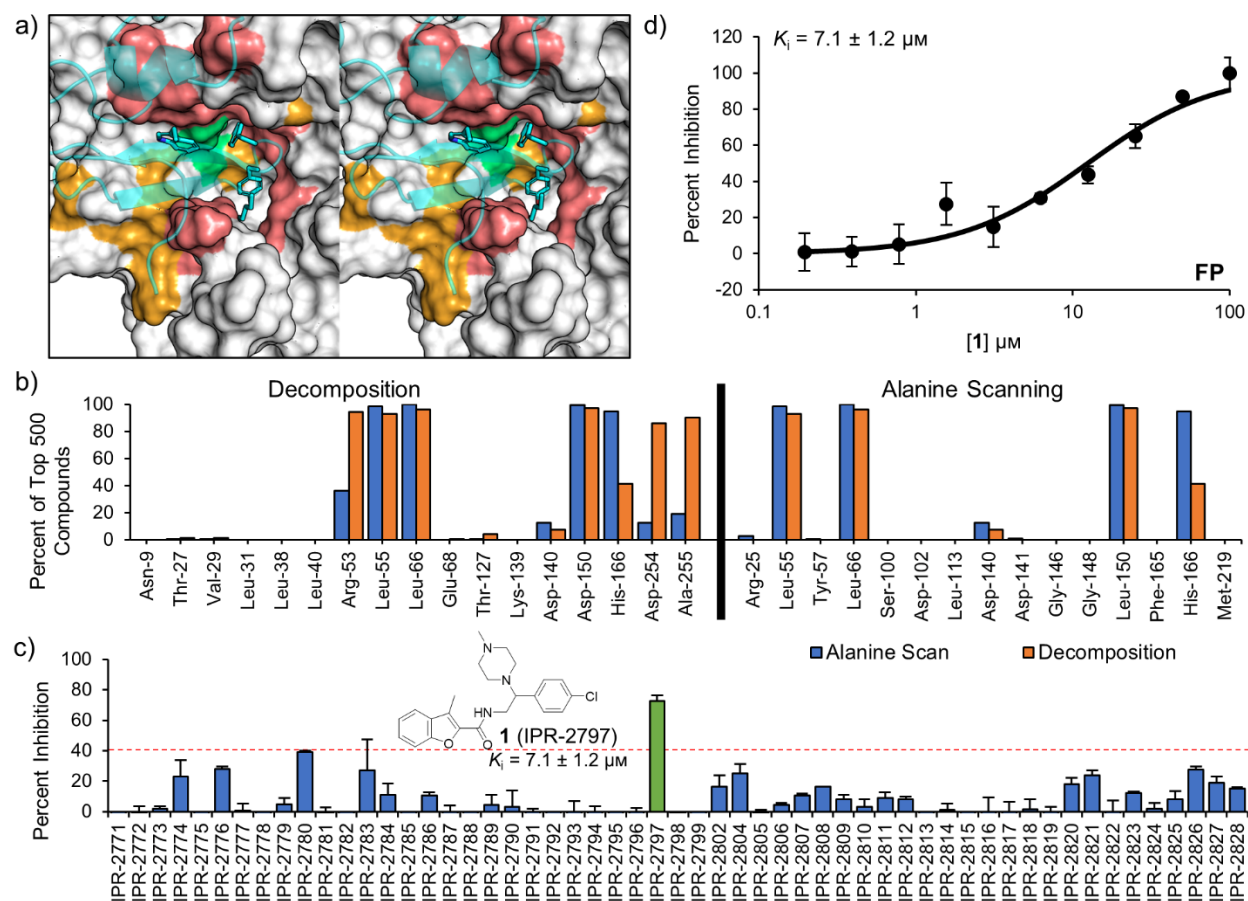


Figure 2

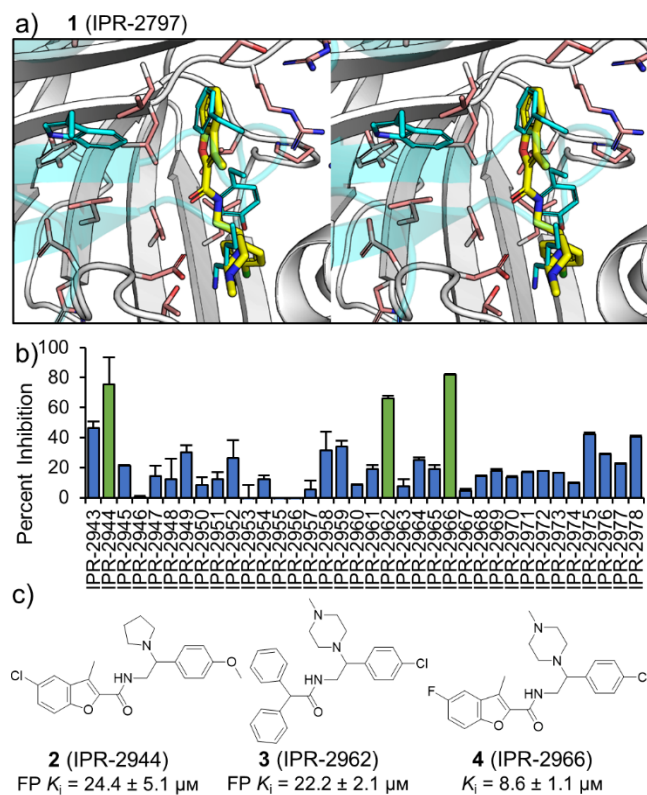


Figure 3

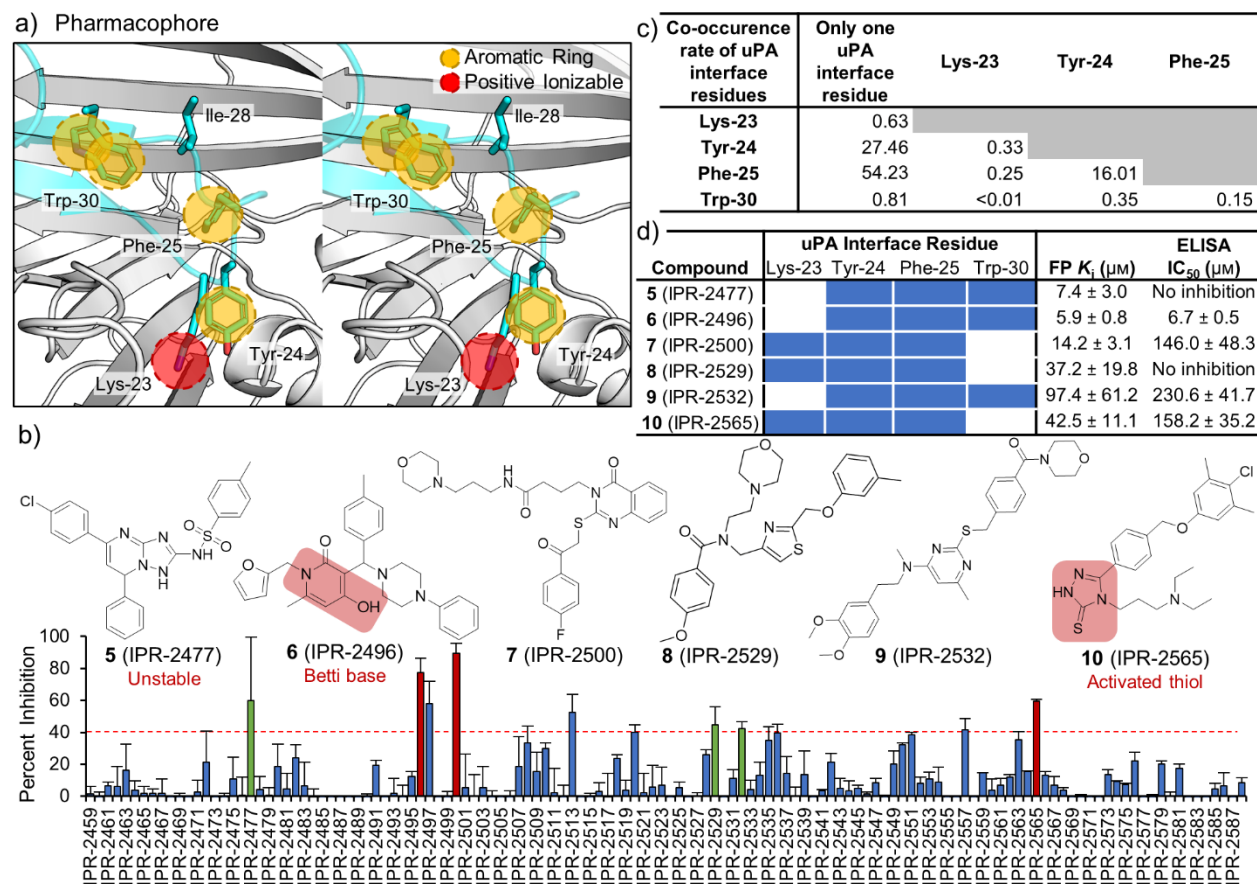


Figure 4

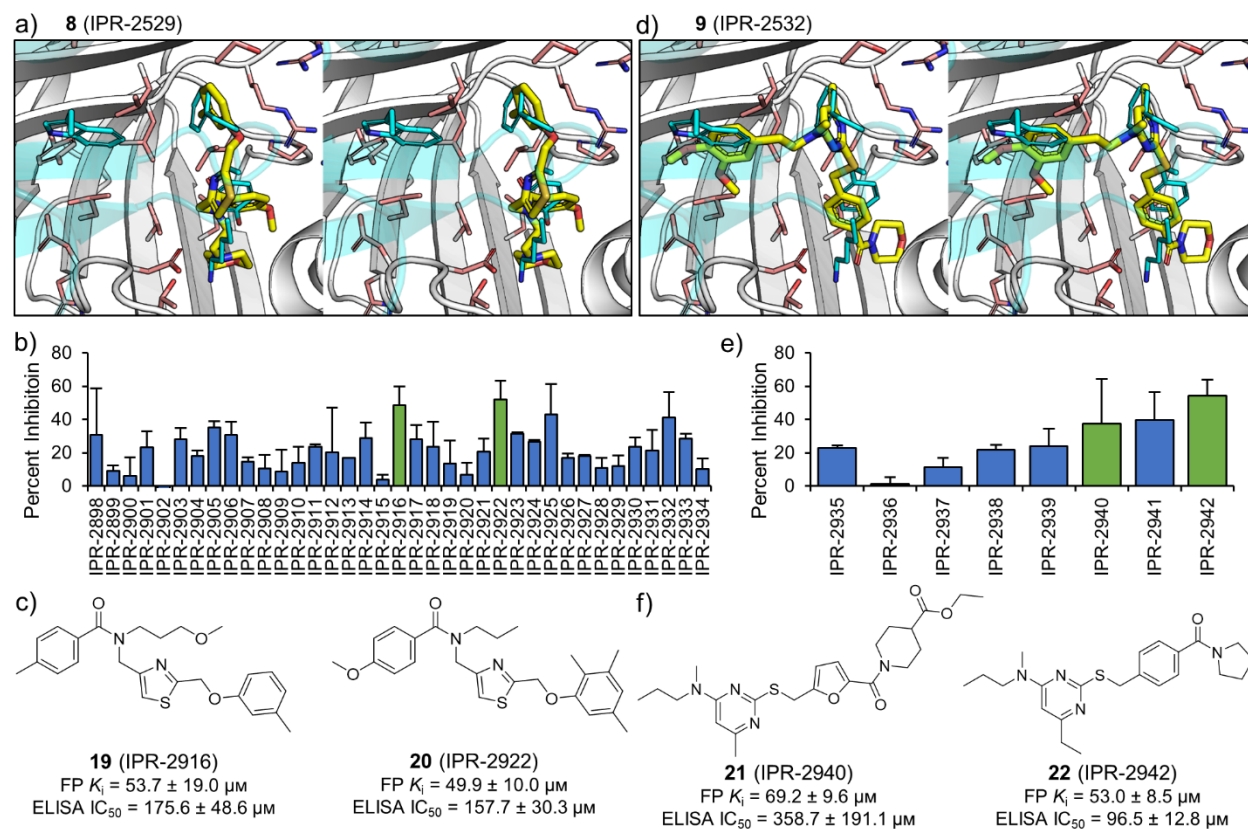


Figure 5

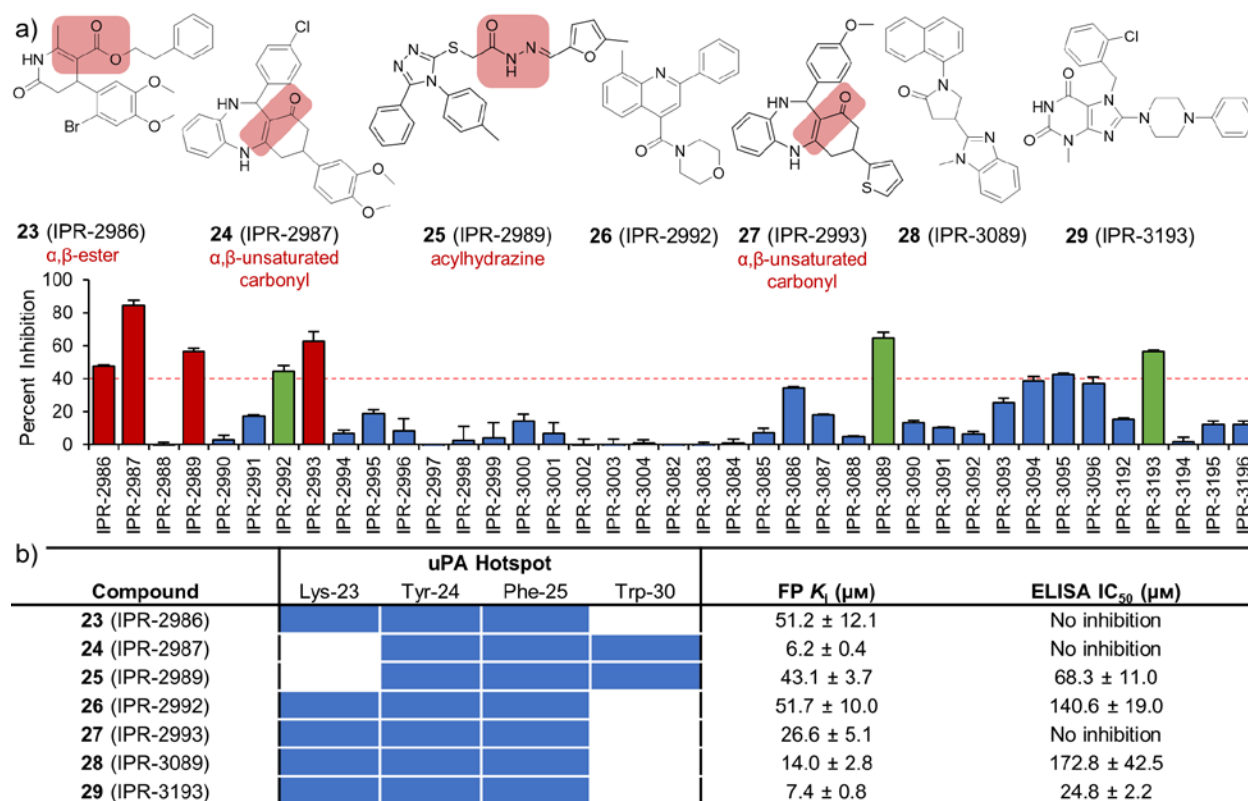


Figure 6

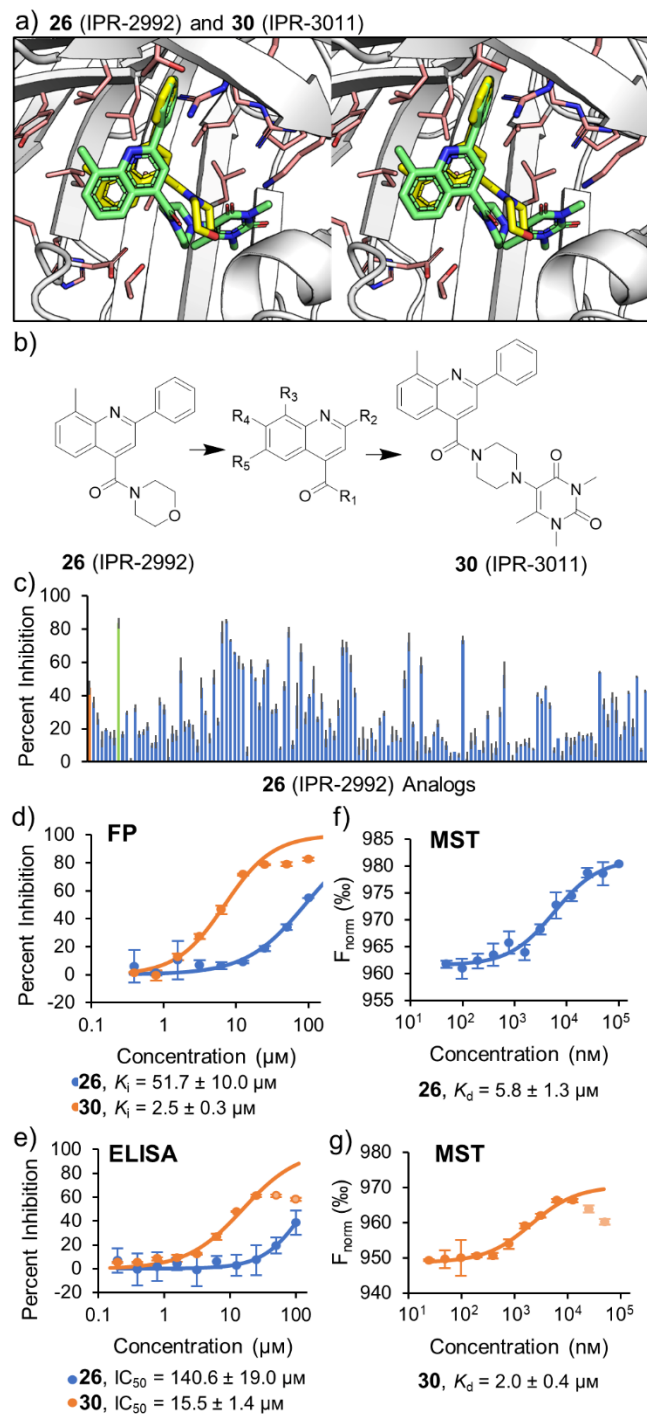


Figure 7

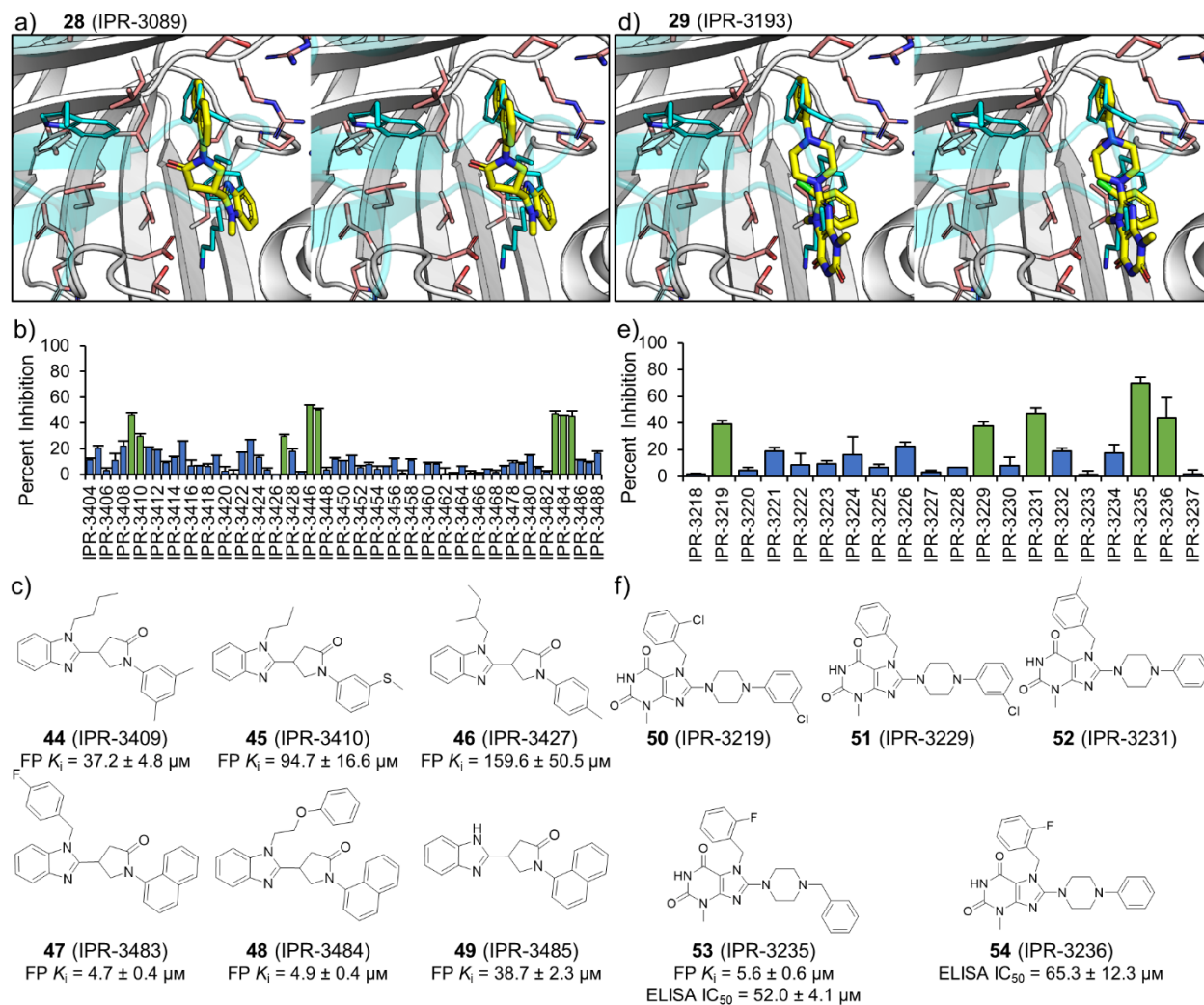


Figure 8